MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

ON THE SAMPLE MEAN AND VARIANCE OF A LONG MEMORY PROCESS

by

Donald B. Percival

TECHNICAL REPORT No. 69

September, 1985

Department of Statistics, GN-22

University of Washington

Seattle, Washington 98195

USA

On the Sample Mean and Variance of a Long Memory Process

Donald B. Percival

Department of Statistics
University of Washington

ABSTRACT

Some properties of the sample mean and sample variance of a long memory process are described. It is shown that for a particular class of long memory processes the asymptotic relative efficiency of the k-decimated sample mean $\bar{x}_n(k)$ (formed by taking the mean of every k-th observation) with respect to the sample mean $\bar{x}_n$ of all n observations is 1, but that the deficiency (as defined by Hodges and Lehmann) of $\bar{x}_n(k)$ with respect to $\bar{x}_n$ is infinite. It is also shown that the sample variance $\hat{\sigma}_n^2$ of a long memory process can be badly biased toward 0: for any integer N and every $\epsilon > 0$ there exists a long memory process with variance σ^2 such that $E\hat{\sigma}_n^2 < \epsilon\sigma^2$ for all sample sizes $n \leq N$. These properties of the sample mean and variance are not shared by "ordinary" stationary processes such as ARMA processes.

1. Introduction and Summary

Let $\{x_t : t = 0, \pm 1, \pm 2, \cdots\}$ represent a wide-sense (or covariance) stationary process that is zero mean, real-valued, discrete time, and purely continuous. Let

$$C_x(\tau) = E x_t x_{t+\tau}, \quad \tau = 0, \pm 1, \pm 2, \cdots,$$

represent the autocovariance function (a.c.f.) of the process, and let S_x represent its spectral density function (s.d.f.) defined on the interval $[-\pi, \pi]$. In the classic treatment of stationary processes that forms the basis for time series analysis, it is usually assumed that $C_x(\tau) \to 0$ "fairly rapidly" as $\tau \to \infty$ or that S_x is bounded on $[-\pi, \pi]$. In particular, the regularity condition $\sum |C_x(\tau)| < \infty$ (which implies that S_x is continuous everywhere and hence bounded) is often seen as a hypothesis in various theorems (see, for example, the discussion in section 1.3 of Brillinger[5]). This condition is satisfied by many useful models for time series, including all stationary and invertible autoregressive, moving average (ARMA) processes with a finite number of parameters (see, for example, Box and Jenkins[4]).

In spite of the pervasiveness of these regularity conditions in time series analysis, there are many time series in nature for which the appropriateness of these conditions is questionable. Empirically such series have two characteristics: the observed correlation between data points that are τ units apart does not decrease at the rapid exponential rate of ARMA models as τ increases; and estimates of the s.d.f. show that it peaks sharply at the origin. Examples include the difference in apparent time as generated by two different atomic clocks[20]; fluctuations in the earth's rate of rotation[17]; the annual minimum flood levels on the Nile River[16]; density fluctuations of sand particles passing through an hour glass[23]; density fluctuations of traffic on an expressway[18]; pitch fluctuations in Western music[25]; fluctuations in the arrival times of pulses from pulsars[7]; voltage fluctuations across cell membranes[14]; deviations in resistance of a 56K ohm India ink resistor[26]; and a wide variety of economic time series[9].

To provide models these many time series, we consider stationary processes in this report such that $\sum |C_x(\tau)| = \infty$ or that S_x has a singularity at the origin. Such processes have been termed long memory processes (l.m.p.'s) in the literature. In section 2 below, we discuss some precise definitions of a l.m.p. and describe an important example of a l.m.p., namely, the power law l.m.p. of order $-1 < \alpha < 0$, for which

$$\lim_{\omega \to 0} \frac{S(\omega)}{|\omega|^\alpha} = h$$

for some constant h. A specific example of a power law l.m.p. that is analytically tractable is the fractional difference process (f.d.p.) introduced by Granger and Joyeux[10].

In section 3, we show that, if $\{x_t\}$ is a power law l.m.p. of order α with mean μ_x, the sample mean $\bar{x}_n$ converges a.s. to μ_x and

$$var \ \bar{x}_n = O(\frac{1}{n^{1+\alpha}}).$$

Next we show that, if a power law l.m.p. has a s.d.f. such that the limits $S(\omega \pm)$ exist for all ω in $(-\pi, \pi)$, then, for any integer $k \geq 2$, its k-decimated sample mean (formed by taking the mean of every k-th observation) is asymptotically as efficient as the ordinary sample mean. As a converse, we prove that, if a stationary process has a summable a.c.f., there exists a k_0 such that the asymptotic relative efficiency of the k-decimated sample mean with respect to

the ordinary sample mean is strictly less than 1 for all $k \geq k_0$. We next show that for a f.d.p. the deficiency (as defined by Hodges and Lehmann) of the k-decimated sample mean with respect to the sample mean is infinite. In passing, we give some necessary and sufficient conditions on the a.c.f. of a stationary process such that

$$var \; \overline{y}_n < var \; \overline{y}_{n-1} \; .$$

In section 4, we show that the sample variance s_n^2 of a f.d.p. can be a seriously biased estimator of the process variance. In particular, we show that, for every sample size N and every $\varepsilon > 0$, there exists a long memory process with variance σ^2 such that $Es_n^2 < \varepsilon\sigma^2$ for all $n \leq N$. In practice, this result means that the process variance can be a poor measure of variability of a l.m.p.

2. Characterization and Models for Long Memory Processes

The processes of main interest in this report are such that $\sum |C_x(\tau)| = \infty$; accordingly, we make the following

Definition 1: A wide-sense stationary process $\{x_t\}$ with a.c.f. C_x is called a long memory process (l.m.p.) in the covariance sense if $\sum |C_x(\tau)| = \infty$.

The qualifier "in the covariance sense" is needed since other characterizations of a l.m.p. are possible and are not necessarily equivalent to the covariance definition. In addition to this covariance characterization, Parzen[19] proposes definitions in terms of the s.d.f. and four other criteria. For what follows, a characterization in terms of the s.d.f. is convenient, so we first review under what conditions 1) a stationary process has a s.d.f. and 2) a given function is a s.d.f. for some stationary process.

By the Wiener-Khinchin theorem (see, for example, volume 1, p.222 of Priestley[22]), a necessary and sufficient condition that C_x be the a.c.f. for some stationary process $\{x_t\}$ is that there exists a function F_x (the spectral distribution function) defined on $[-\pi,\pi]$ such that $F_x(-\pi) = 0$; $0 < F_x(\pi) < \infty$; F_x is non-decreasing on $[-\pi,\pi]$; and

$$C_x(\tau) = \int_{-\pi}^{\pi} e^{i\omega\tau} dF_x(\omega) \; .$$

If F_x is absolutely continuous with respect to Lebesgue measure, then

$$F_x(\lambda) = \int_{-\pi}^{\lambda} S_x(\omega) d\omega \; ,$$

where S_x is called a s.d.f. for the process. S_x is necessarily non-negative on $[-\pi,\pi]$; even; and unique up to sets of Lebesgue measure zero since, if $S_x^{'}$ differs only on a set of measure zero from S_x, $S_x^{'}$ is also a s.d.f. for $\{x_t\}$. Moreover,

$$C_x(\tau) = \int_{-\pi}^{\pi} e^{i\omega\tau} S_x(\omega) d\omega = 2\int_{0}^{\pi} \cos\omega\tau \; S_x(\omega) d\omega \; .$$

If S is to be a s.d.f. for some stationary process, it follows from the Weiner-Khinchin theorem that it is necessary and sufficient that S be even, non-negative on $[-\pi,\pi]$ a.e., positive on a set of positive measure and finitely integrable.

Definition 2: A wide-sense stationary process $\{x_t\}$ with s.d.f. S_x is called a l.m.p. in the s.d.f. sense if the spectral ratio

$$\frac{ess \; sup \; S_x(\omega)}{ess \; inf \; S_x(\omega)} = \infty \; .$$

Both definitions 1 and 2 are essentially due to Parzen[19]; however, he does not discuss the relationship between his covariance and s.d.f. characterizations, so the following lemma is of some interest.

Lemma 3: A l.m.p. in the covariance sense need not be a l.m.p. in the s.d.f. sense. The converse is also true.

Proof: Let $S_x(\omega) = 1 + I_{[0,b]}(|\omega|)$, where I_A is the indicator function for the set A and $0 < b < \pi$. Then the spectral ratio is 2, but $C_x(\tau) = (2\sin\tau b)/\tau$, so $\sum |C_x(\tau)| = \infty$. Conversely, let $S_x(\omega) = 4\sin^2\frac{\omega}{2}$. Then the spectral ratio is infinite, but for $\tau \geq 2$ $C_x(\tau) = 0$ so $\sum |C_x(\tau)| < \infty$. $\square$

Of course, there exist processes that are l.m.p.'s in both senses and also processes that are not l.m.p.'s in either sense. The fractional difference processes that are described below are examples of the former; stationary and invertible ARMA processes with a finite number of parameters are examples of the latter.

The proof of lemma 3 suggests that definition 2 is not a very good one for our purposes: a process with a s.d.f. of $4\sin^2\frac{\omega}{2}$ is just the first difference of a white noise process and intuitively should not qualify as a l.m.p. An examination of the time series that are mentioned in section 1 suggests a narrowing of the definition of a l.m.p. in the s.d.f. sense. Estimates of the s.d.f. of those time series show a concentration of power at low Fourier frequencies with a smooth tapering off toward higher frequencies. Accordingly, we make the following

Definition 4: A wide-sense stationary process $\{x_t\}$ with s.d.f. S_x is called a l.m.p. in the *restricted* s.d.f. sense if S_x satisfies two conditions: 1) it is bounded above on $[\delta,\pi]$ for every $\delta > 0$ and 2) the limit $S_x(0+) = \infty$.

We first need to show that the class of l.m.p.'s in the restricted s.d.f. sense is not empty.

Lemma 5: Let the function S be integrable, even, non-negative, and satisfy the two restrictions on a s.d.f. of a l.m.p. in the restricted s.d.f. sense. Suppose that S is such that

$$\lim_{\omega \to 0} \frac{S(\omega)}{|\omega|^\alpha} = h \qquad (6)$$

for some $0 > \alpha > -1$ and $0 < h < \infty$. Then S is in fact a s.d.f. for some stationary process $\{x_t\}$.

Proof: It follows from the Wiener-Khinchin theorem that we need only show that S integrates finitely over $[-\pi,\pi]$ to a positive value. Since $S(0+) = S(0-) = \infty$, there exists a $\delta > 0$ such that, if $|\omega| < \delta$, $S(\omega) \geq 1$. Then

$$\int_{-\pi}^{\pi} S(\omega)\, d\omega \geq \int_{-\delta}^{\delta} S(\omega)\, d\omega \geq 2\delta,$$

so the integral of S is positive. To see that it integrates finitely, pick $\varepsilon > 0$ and find a $\delta > 0$ such that

$$|S(\omega) - h|\omega|^\alpha| < \varepsilon|\omega|^\alpha$$

when $|\omega| < \delta$. We have

$$\int_{-\pi}^{\pi} S(\omega)\, d\omega \leq 2\int_{0}^{\pi} |S(\omega) - h\omega^\alpha| + h\omega^\alpha\, d\omega$$

$$\leq (2\varepsilon + 4h)\frac{\pi^{1+\alpha}}{1+\alpha} + 2(\pi - \delta)B < \infty,$$

where $B \equiv \sup_{\omega \in [\delta, \pi]} S(\omega)$. $\square$

Obviously a l.m.p. in the restricted s.d.f. sense is also a l.m.p. in the s.d.f. sense. The relationship between the restricted s.d.f. and the covariance characterization is given by

Lemma 7: A l.m.p. in the restricted s.d.f. sense is also a l.m.p. in the covariance sense.

Proof: Let C_x and S_x be the a.c.f. and a s.d.f., respectively, for $\{x_t\}$, a l.m.p. in the restricted s.d.f. sense. We use the same conventions for Fourier series coefficients as Titchmarch[24] in what follows. Since $C_x(\tau)/\pi$ is simply the τ-th Fourier coefficient of S_x,

$$s \equiv \frac{1}{2\pi} \{C_x(0) + 2\sum_{\tau=1}^{\infty} C_x(\tau)\}$$

is the Fourier series of S_x evaluated at $\omega = 0$. Let

$$s_n \equiv \frac{1}{2\pi} \{C_x(0) + 2\sum_{\tau=1}^{n} C_x(\tau)\}$$

be the n-th partial sum associated with s. Now the n-th arithmetic mean or Cesàro sum of $\{s_n\}$ is defined by

$$\sigma_n = \frac{1}{n}\sum_{k=0}^{n-1} s_k = \frac{1}{2n\pi}\int_{-\pi}^{\pi} \frac{\sin^2\frac{n\omega}{2}}{\sin^2\frac{\omega}{2}} S_x(\omega)d\omega$$

by a standard argument in Fourier analysis (see, for example, p.412 of Titchmarch[24]). We now use part of Fejér's theorem in the form given by p.89 of Zygmund[27]: since the limits $S_x(0\pm)$ exist and are both ∞, $\sigma_n \to \infty$. Next define

$$s_n^* = \frac{1}{2\pi}(C_x(0) + 2\sum_{\tau=1}^{n}|C_x(\tau)|)$$

and

$$\sigma_n^* = \frac{1}{n}\sum_{k=0}^{n-1} s_k^*.$$

Since $s_n^* \geq s_n$, $\sigma_n^* \geq \sigma_n$ and $\sigma_n^* \to \infty$ also. Assume for the moment that s_n^* does not diverge to ∞. Since s_n^* is a non-decreasing sequence, it has a finite limit, say, s^*. Pick N so large that $|s_n^* - s^*| < \varepsilon$ for all $n \geq N$. For $n > N$

$$|\sigma_n^* - s^*| \leq \frac{1}{n+1}\{\sum_{k=0}^{N-1}|s_k^* - s^*| + \sum_{k=N}^{n-1}|s_k^* - s^*|\} \leq \frac{K}{n+1} + \varepsilon,$$

where K is a constant independent of n. Thus σ_n^* converges to s^* also, which is a contradiction. We conclude that $s_n^* \to \infty$ which readily implies the lemma. $\square$

The next lemma concludes our discussion on the relationship among s.d.f., restricted s.d.f. and covariance sense l.m.p.'s.

Lemma 8: There exist processes that are l.m.p.'s in both the s.d.f. and covariance sense but not in the restricted s.d.f. sense.

Proof: Let S_x be a s.d.f. for a l.m.p. in the restricted s.d.f. sense, and define $S_y(\omega) = S_x(\pi - |\omega|)$. S_y is clearly not a s.d.f. for any l.m.p. in the restricted s.d.f. sense, but is a s.d.f. for some l.m.p. $\{y_t\}$ in the s.d.f. sense. It follows easily that

$$C_y(\tau) = (-1)^{\tau} C_x(\tau) .$$

Lemma 7 shows that $\sum |C_x(\tau)| = \infty$, so $\sum |C_y(\tau)| = \infty$ also and the lemma follows. $\square$

Figure 1 summarizes lemmas 3, 7, and 8 schematically. Each point in this figure represents a l.m.p. in whatever senses the loops that surround it are labelled.
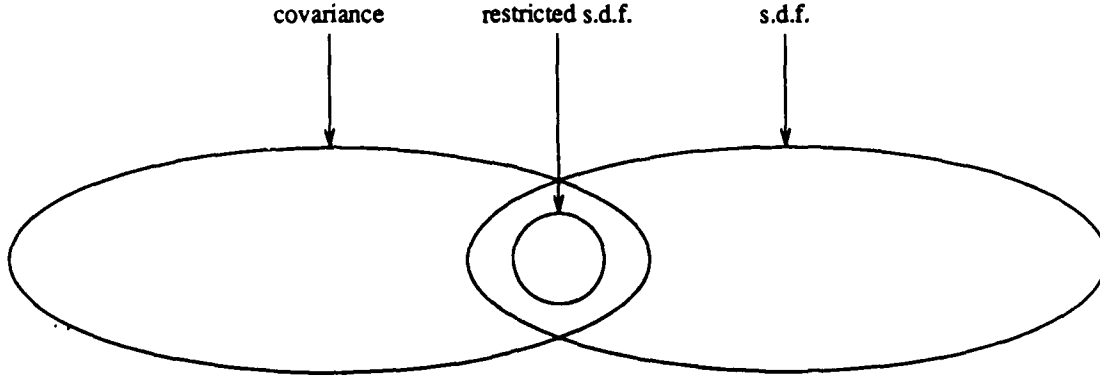

Figure 1: Relationship between Various L.M.P.'s

The rate at which the s.d.f. for a l.m.p. in the restricted s.d.f. sense diverges to infinity as $|\omega| \to 0$ is important in some of the results to follow. Equation (6) gives one particular rate, namely that S is approximately proportional to $|\omega|^{\alpha}$ near the origin.

Definition 9: The function S is said to obey a power law in the limit (as $\omega \to 0$) of order α, where α is any real number, if

$$\lim_{\omega \to 0} \frac{S(\omega)}{|\omega|^{\alpha}} = h$$

for some $0 < h < \infty$.

Definition 10: A wide-sense stationary process $\{x_t\}$ with a s.d.f. S_x is called a power law l.m.p. of order α, where $-1 < \alpha < 0$, if S_x 1) obeys a power law in the limit of order α and 2) is bounded above on every interval $[\delta, \pi]$ with $\delta > 0$.

It is obvious that a power law l.m.p. is just a special type of l.m.p. in the restricted s.d.f. sense. The basis for definition 10 is that empirical s.d.f.'s for the time series mentioned in section 1 often seem to have a linear appearance near zero Fourier frequency when they are plotted on a log-log scale. The slope of the line is given by the exponent α. For example, Mohr[16] finds that $\alpha = -.67$ gives a good fit for data on the yearly minimum water levels for the Nile River.

Let us now consider an example of a power law l.m.p. that is analytically tractable.

Definition 11: A wide-sense stationary process $\{x_t\}$ is called a fractional difference process (f.d.p.) of order $\alpha > -1$ if it has a s.d.f. given by

$$S_x(\omega) = h \left(4\sin^2\frac{\omega}{2}\right)^{\frac{\alpha}{2}} \tag{12}$$

for some $0 < h < \infty$. (That S_x obeys a power law in the limit of order α follows from an application of l'Hospital's rule.)

The name "fractional difference" is derived from the following considerations. Suppose that $\{y_t\}$ is a white noise process with a s.d.f. $S_y(\omega) = h$. Let $\{z_t\}$ represent the process formed by taking the n-th finite differenc of $\{y_t\}$. Then $\{z_t\}$ has a s.d.f. given by

$$S_z(\omega) = h \, (4\sin^2 \frac{\omega}{2})^n \, .$$

Thus S_x in equation (12) may be regarded as a s.d.f. for a "fractionally differenced" white noise process. (The order of the f.d.p. in definition 11 is defined as α instead of $\alpha/2$ to remind us that $S_x(\omega) \approx h \, |\omega|^{\alpha}$ for ω close to 0.)

This model is due to Granger and Joyeux[10], to whom the reader is referred for more details. They show that the a.c.f. of this process is given by

$$C_x(\tau) = -2h \sin \frac{\pi\alpha}{2} \, \Gamma(1+\alpha) \frac{\Gamma(\tau - \frac{\alpha}{2})}{\Gamma(\tau + 1 + \frac{\alpha}{2})} \equiv h_\alpha \frac{\Gamma(\tau - \frac{\alpha}{2})}{\Gamma(\tau + 1 + \frac{\alpha}{2})} \tag{13}$$

for $\tau \geq 0$ when $-1 < \alpha < 2$ and $\alpha \neq 0$. We note that it is easy to compute C_x step by step since

$$C_x(\tau) = C_x(\tau - 1) \frac{2\tau - \alpha - 2}{2\tau + \alpha} \tag{14}$$

for $\tau = 1, 2, \cdots$. It is sometimes useful to have a good approximation to C_x, so the following lemma gives an error analysis for an approximation due to Granger and Joyeux.

Lemma 15: $C_x(\tau) = \frac{h_\alpha}{\tau^{1+\alpha}} \, (1 + O(\frac{1}{\tau^2}))$ for $\tau \geq 1$; moreover, if $-1 < \alpha < 0$, the bounding constant for the O term may be taken as .35.

Proof: The proof of this lemma is messy and uninteresting. The interested reader is referred to Percival[21] for details. $\square$

An immediate corollary is that $\sum |C_x(\tau)| = \infty$ if $-1 < \alpha < 0$ (as must be true by lemma 7 also) and $\sum |C_x(\tau)| < \infty$ if $\alpha \geq 0$.

A f.d.p. has a number of mathematical advantages when a specific model is needed for a l.m.p. since its s.d.f. has a simple mathematical form given by equation (12). In contrast, the discrete fractional Gaussian process (d.f.g.p.) that Mandelbrot[15] has introduced does not have this property. A d.f.g.p. $\{x_t\}$ is defined as a stationary Gaussian process that has a a.c.f. given by

$$C_x(\tau) = \frac{C_x(0)}{2} \{|\tau + 1|^{1-\alpha} - 2|\tau|^{1-\alpha} + |\tau - 1|^{1-\alpha}\}$$

for $|\tau| \geq 1$, $C_x(0) > 0$, and $-1 < \alpha \leq 1$. For large positive τ,

$$C_x(\tau) \approx \frac{C_x(0)\alpha(\alpha - 1)}{2\tau^{1+\alpha}} \, ,$$

so the a.c.f.'s for a d.f.g.p. and a f.d.p. have approximately the same structure for large τ. It is also true that a d.f.g.p. has a s.d.f. that obeys a power law in the limit of order α, but unfortunately this s.d.f. cannot be expressed in a closed form. Moreover, the simple, yet accurate, approximation to the exact a.c.f. of a f.d.p. given by lemma 1.23 is often easier to work with than the rather cumbersome exact expression for the a.c.f. of a d.f.g.p. For this reason, we prefer f.d.p.'s to d.f.g.p.'s.

- 8 -

1.3 The Sample Mean of a Power Law L.M.P.

Let $\{x_t\}$ represent a power law l.m.p. of order $-1 < \alpha < 0$ with s.d.f. S_x. Let $Ex_t = \mu_x$. Suppose that we are given a finite portion of a realization of this process, say, $x_1, x_2, \cdots, x_n$, and that we want to estimate μ_x. The natural estimator to consider is the sample mean

$$\bar{x}_n \equiv \frac{1}{n}\sum_{t=1}^{n} x_t .$$

Strong consistency of this estimator is shown in lemma 17, for which we need the following simple result:

Lemma 16: Let $\{x_t\}$ be a power law l.m.p. of order $-1 < \alpha < 0$ with s.d.f. S_x. Then 1) $S_x(\omega) = O(|\omega|^\alpha)$ and 2) $S_x(\omega) - h|\omega|^\alpha = o(|\omega|^\alpha)$ as $\omega \to 0$.
Proof: Suppose first that $-1 < \alpha < 0$. Pick any $\varepsilon > 0$. There exists a $\delta \in (0, \pi)$ such that

$$|S_x(\omega) - h\omega^\alpha| < \varepsilon\omega^\alpha$$

for all $0 < \omega < \delta$, which yields statement 2) of the lemma. The above implies that

$$S_x(\omega) < (h + \varepsilon)\omega^\alpha$$

for $0 < \omega < \delta$. Moreover, since S_x is bounded above for $\omega \in [\delta, \pi]$, it follows that

$$S_x(\omega) \leq K|\omega|^\alpha$$

on $[-\pi, \pi]$ for some finite constant K and hence that statement 1) holds. $\square$

Lemma 17: Let $\{x_t\}$, μ_x, and $\bar{x}_n$ be as specified above. Then $\bar{x}_n$ converges to μ_x a.s.
Proof: We use the following result, which is a direct consequence of corollary 3, p.207, of Hannan[11]: if $\{x_t\}$ is a wide-sense stationary process with s.d.f. S_x and if

$$\int_0^\delta S_x(\omega)\, d\omega = O(\delta^\beta)$$

as $\delta \to 0$ for some $\beta > 0$, then $\bar{x}_n$ converges a.s. to μ_x. By lemma 16, $S_x(\omega) \leq K\omega^\alpha$ for $\omega > 0$ and some $0 < K < \infty$. Thus

$$\int_0^\delta S_x(\omega)\, d\omega \leq K\int_0^\delta \omega^\alpha\, d\omega = \frac{K\delta^{1+\alpha}}{1+\alpha} = O(\delta^\beta)$$

for $\beta = 1 + \alpha > 0$ as $\delta \to 0$. $\square$

We want to determine at what rate $var\ \bar{x}_n$ decreases to 0 as $n \to \infty$. Unfortunately, since S_x is discontinuous at 0, we cannot quote the standard result (theorem 6.12, p.232, of Fuller[8]) that

$$var\ \bar{x}_n \approx \frac{2\pi S_x(0)}{n}$$

for large n, which in fact is incorrect. The next lemma shows that the rate of decrease is related to the exponent α.

Lemma 18: Let $\{x_t\}$ be as in lemma 17. Then

$$var\ \bar{x}_n = O(\frac{1}{n^{1+\alpha}}) .$$

Proof: By lemma 16, $S_x(\omega) \leq K|\omega|^\alpha$ on $[-\pi, \pi]$ for some constant K. By theorem 8.23, p.444, of Anderson[3], we have

$$var \ \bar{x}_n = \frac{1}{n} \int_{-\pi}^{\pi} \frac{\sin^2 \frac{n\omega}{2}}{n \sin^2 \frac{\omega}{2}} S_x(\omega) \, d\omega$$

$$\leq 2K \int_0^{\frac{1}{n}} \omega^\alpha d\omega + \frac{2K}{n^{1+\alpha}} \int_{\frac{1}{n}}^{\pi} \frac{\sin^2 \frac{n\omega}{2}}{n \sin^2 \frac{\omega}{2}} d\omega$$

$$\leq \frac{2K}{(1+\alpha)n^{1+\alpha}} + \frac{2K}{n^{1+\alpha}} \int_0^{\pi} \frac{\sin^2 \frac{n\omega}{2}}{n \sin^2 \frac{\omega}{2}} d\omega$$

$$= \frac{2K\{1+\pi(1+\alpha)\}}{(1+\alpha)n^{1+\alpha}} \ ,$$

where we have made use of the following properties of the Fejér kernel:

$$0 \leq \frac{\sin^2 \frac{n\omega}{2}}{n \sin^2 \frac{\omega}{2}} \leq n$$

(equation (36), p.465 of Anderson[3]) and

$$\int_0^{\pi} \frac{\sin^2 \frac{n\omega}{2}}{n \sin^2 \frac{\omega}{2}} d\omega = \pi$$

(equations (9) and (11), p.461 of Anderson[3]). $\square$

Let us now specialize to the case where $\{x_t\}$ is a f.d.p. of order $-1 < \alpha < 0$. The s.d.f. S_x and a.c.f. C_x for $\{x_t\}$ are given by equations (12) and (13) respectively. We now give a more precise rate of decrease for $var \ \bar{x}_n$ in the following lemma. For use below the lemma actually gives the rate of decrease of the variance of a k-decimated mean, defined by

$$\bar{x}_n(k) = \frac{1}{m} \sum_{i=1}^{m} x_{ki} \ ,$$

where m is the largest integer that does not exceed n/k. A k-decimated mean thus uses only every k-th data value. Since $\bar{x}_n(1) = \bar{x}_n$, the case $k = 1$ gives the rate of decrease for $var \ \bar{x}_n$.

Lemma 19: $var \ \bar{x}_n(k) = \dfrac{2h_\alpha}{n^{1+\alpha}\alpha(\alpha-1)} + O(\dfrac{1}{n})$ for a f.d.p. with $-1 < \alpha < 0$; moreover, the O term is less than

$$C_x(0)\{\frac{A}{n} + \frac{B}{(1+\alpha)n^2}\},$$

in magnitude, where A and B are constants independent of α.

Proof: The a.c.f. for $\{x_{ki}\}$ is just $C_x(k\tau)$, $\tau = 0, 1, \cdots$. By a simple extension of a well-known result (see, for example, the proof of corollary 6.1.1.2, p.232, of Fuller[8]),

$$var\ \bar{x}_n(k) = \frac{1}{m}\{C_x(0) + 2\sum_{\tau=1}^{m}(1 - \frac{\tau}{m})C_x(k\tau)\}\ . \tag{20}$$

By lemma 15, the summation above is equal to

$$\frac{h_\alpha}{k^{1+\alpha}}\{\sum_{\tau=1}^{m}(1-\frac{\tau}{m})\frac{1}{\tau^{1+\alpha}} + \frac{1}{k^2}\sum_{\tau=1}^{m}(1-\frac{\tau}{m})O(\frac{1}{\tau^{3+\alpha}})\}\ . \tag{21}$$

Define

$$f(t) = (1 - \frac{t}{m})t^\beta\ .$$

A simple argument shows that the values that any given derivative of f assumes on the interval $(1, m)$ all have the same sign if $\beta < 0$. In this case we may use the Euler-Maclaurin summation formula (see, for example, section 5.8 of Hildebrand[12])

$$\sum_{\tau=1}^{m}f(\tau) = \int_1^m f(t)\,dt + \frac{1}{2}\{f(1) + f(m)\} + E(m;\beta)$$

and know that

$$E(m;\beta) \le \frac{1}{12}|f'(m) - f'(1)|\ ,$$

where f' denotes the first derivative of f; and $E(m;\beta)$ has the same sign as the quantity inside the absolute value signs. Computations show that, if $\beta \ne -1$ or -2,

$$\sum_{\tau=1}^{m}f(\tau) = \frac{m^{\beta+1}}{(\beta+1)(\beta+2)} - \frac{1}{\beta+1} + \frac{1}{(\beta+2)m} + \frac{1}{2} - \frac{1}{2m} + E(m;\beta)\ ,$$

where

$$0 \le E(m;\beta) \le \frac{1}{12}\{-\frac{1}{m^{1-\beta}} - \beta + \frac{\beta+1}{m}\}\ .$$

If we let $\beta = -1-\alpha$, then $-1 < \beta < 0$ and we may use the above to show that the first term in the brackets in (21) is equal to

$$\frac{m^{-\alpha}}{\alpha(\alpha-1)} + \frac{1}{\alpha} + \frac{1}{(1-\alpha)m} + \frac{1}{2} - \frac{1}{2m} + E(m;-1-\alpha)\ ,$$

with $E(m;-1-\alpha) = O(1)$. By the second part of lemma 15, the second term in the brackets, which we denote as $R(m)$, is such that

$$|R(m)| < \frac{.35}{k^2}\sum_{\tau=1}^{m}(1 - \frac{\tau}{m})\frac{1}{\tau^{3+\alpha}}\ .$$

If we now let $\beta = -3-\alpha$, then $-3 < \beta < -2$ and it follows from the above that

$$|R(m)| < \frac{.35}{k^2}\{\frac{1}{m^{2+\alpha}(1+\alpha)(2+\alpha)} + \frac{1}{2+\alpha} - \frac{1}{m(1+\alpha)}$$
$$+ \frac{1}{2} - \frac{1}{2m} + E(m;-3-\alpha)\}\ .$$

Since $E(m;-3-\alpha) = O(1)$, $R(m) = O(1)$ also. By combining the above results we have

$$var\ \bar{x}_n(k) = \frac{C_x(0)}{m} + 2h_\alpha\{\frac{1}{(mk)^{1+\alpha}\alpha(\alpha-1)} + \frac{1}{m^2k^{1+\alpha}}[\frac{1}{1-\alpha} - \frac{1}{2}]$$

$$+ \frac{1}{mk^{1+\alpha}} [\frac{1}{\alpha} + \frac{1}{2} + E(m;-1-\alpha) + R(m)]\} \tag{22}$$

$$= \frac{2h_\alpha}{(mk)^{1+\alpha}\alpha(\alpha+1)} + O(\frac{1}{m})$$

Since $mk = n + O(1)$, an additional easy argument yields the first part of the lemma.

To simplify notation, suppose $k = 1$ and let $E \equiv E(n;-1-\alpha)$ and $R \equiv R(n)$. Equation (13) and the above show that the absolute value of the O term in the statement of the first part of this lemma is equal to

$$C_x(0) \left| \frac{1}{n} + \frac{2\Gamma(1+\frac{\alpha}{2})}{\Gamma(-\frac{\alpha}{2})} \{ \frac{1}{n^2} [\frac{1}{1-\alpha} - \frac{1}{2}] + \frac{1}{n} [\frac{1}{\alpha} + \frac{1}{2} + E + R] \} \right|$$

$$\leq C_x(0) \{ \frac{1}{n} - \frac{\alpha\Gamma(1+\frac{\alpha}{2})}{n\,\Gamma(1-\frac{\alpha}{2})} [-\frac{1}{\alpha} + 1 + |E| + |R|] \} .$$

Since $|E| < 1$ and

$$|R| < C + \frac{B}{(1+\alpha)n}$$

for finite constants B and C independent of α, it follows that the O term is bounded in absolute value by

$$C_x(0) \{ \frac{A}{n} + \frac{B}{(1+\alpha)n^2} \}$$

for some finite constant A independent of α. A similar argument holds when $k \geq 2$. $\square$

The most interesting aspect of this lemma is that the leading term for $var \; \bar{x}_n(k)$ is independent of k. We now make use of this observation to show an interesting property of a f.d.p. Let

$$e(k;n) \equiv \frac{var \; \bar{x}_n}{var \; \bar{x}_n(k)} \tag{23}$$

be the relative efficiency and, if the limit exists,

$$e(k) \equiv \lim_{n \to \infty} e(k;n)$$

be the asymptotic relative efficiency (a.r.e.) of $\bar{x}_n(k)$ with respect to $\bar{x}_n$.

Corollary 24: For a f.d.p. with $-1 < \alpha < 0$, $e(k) = 1$ for all k.

Proof: This follows immediately from lemma 19. $\square$

Loosely speaking, corollary 24 says that we can discard an arbitrarily large fixed proportion of observations from a f.d.p. of order $-1 < \alpha < 0$ and still produce an estimate of μ_x that is as efficient asymptotically as the sample mean of all the observations. This result indicates how "strong" the memory can be in a l.m.p. For contrast, note that, if $\{x_t\}$ were an uncorrelated process, $e(k) = 1/k$.

As a numerical example of the above results, table 1 gives the exact relative efficiencies, $e(k;n)$ in equation (23), for a f.d.p. with $\alpha = -.5$ and $-.9$ for various values of n and k. For example, for $n = 500$, $\alpha = -.9$, and $k = 50$, $\bar{x}_n(k)$ and $\bar{x}_n$ are averages of 10 and 500 data values, respectively, and yet $e(k;n) = .94$ indicates that $\bar{x}_n(k)$ is almost as efficient as $\bar{x}_n$. For the sake of comparison, the column marked $1/k$ is the value of $e(k;n)$ for an uncorrelated process.

Table 1

Relative Efficiency of $\bar{x}_n(k)$ with Respect to $\bar{x}_n$

k	1/k	$\alpha = -.5$		$\alpha = -.9$	
		$e(k;100)$	$e(k;500)$	$e(k;100)$	$e(k;500)$
2	.5000	.9375	.9711	.9975	.9994
5	.2000	.7627	.8782	.9870	.9970
10	.1000	.5659	.7453	.9653	.9918
20	.0500	.3649	.5636	.9173	.9796
50	.0200	.1721	.3186	.7793	.9392

Theorem 28 below shows that corollary 24 is true for a wider class of power law l.m.p.'s than just f.d.p.'s of order $-1 < \alpha < 0$. Before we can prove that theorem, we need the following result.

Lemma 25: Suppose that $\{y_t\}$ is a power law l.m.p. of order $-1 < \alpha < 0$ with s.d.f. S_y such that the limits $S_y(\frac{2\pi l}{k} \pm)$ exist for $l = 0, \cdots, k-1$ with $k \geq 2$. Then

$$\lim_{n \to \infty} n \{var \ \bar{y}_n(k) - var \ \bar{y}_n\} = \pi \sum_{l=1}^{k-1} \{S_y(\frac{2\pi l}{k}+) + S_y(\frac{2\pi l}{k}-)\} .$$

Proof: The k-decimated mean $\bar{y}_n(k)$ is the sample average of $y_k, y_{2k}, \cdots, y_{mk}$, which we may regard as a sample of a subsequence $\{y_{tk}\}$ of the original process $\{y_t\}$. This subsequence has a s.d.f. that is given by the folding theorem (see, for example, p.388 of Anderson[3]):

$$S_k(\omega) = \frac{1}{k} \sum_{l=0}^{k-1} S_y(\frac{\omega + 2\pi l}{k}) , \qquad (26)$$

where S_y is defined outside of $[-\pi, \pi]$ by periodic extension. The a.c.f. that corresponds to this subsequence is simply $C_k(\tau) = C_y(k\tau)$. In this notation, equation (20) becomes

$$var \ \bar{y}_n(k) = \frac{1}{m} \{C_k(0) + 2\sum_{\tau=1}^{m}(1 - \frac{\tau}{m})C_k(\tau)\}$$

$$= \frac{1}{m^2} \int_{-\pi}^{\pi} \frac{\sin^2 \frac{m\omega}{2}}{\sin^2 \frac{\omega}{2}} S_k(\omega) \, d\omega .$$

(See, for example, p.89 of Zygmund[27].) Define $r = \frac{n}{k} - m$ and note that $0 \leq r < 1$. We then have

$$n \{ var\ \bar{y}_n(k) - var\ \bar{y}_n \}$$

$$= \frac{n}{m^2} \int_{-\pi}^{\pi} \frac{\sin^2 \frac{m\omega}{2}}{\sin^2 \frac{\omega}{2}} S_k(\omega)\, d\omega - \frac{1}{n} \int_{-\pi}^{\pi} \frac{\sin^2 \frac{n\omega}{2}}{\sin^2 \frac{\omega}{2}} S_y(\omega)\, d\omega$$

$$= \frac{1}{2\pi m} \int_{-\pi}^{\pi} \frac{\sin^2 \frac{m\omega}{2}}{\sin^2 \frac{\omega}{2}} f(\omega)\, d\omega + \sum_{i=1}^{4} R_i \ , \tag{27}$$

where

$$f(\omega) = 2\pi \{ k S_k(\omega) - \frac{\sin^2 \frac{\omega}{2}}{k^2 \sin^2 \frac{\omega}{2k}} S_y(\frac{\omega}{k}) \} \ ,$$

$$R_1 = \frac{kr}{m^2} \int_{-\pi}^{\pi} \frac{\sin^2 \frac{m\omega}{2}}{\sin^2 \frac{\omega}{2}} S_k(\omega)\, d\omega \ ,$$

$$R_2 = \frac{r}{mk(m+r)} \int_{-\pi}^{\pi} \frac{\sin^2 \frac{mk\omega}{2}}{\sin^2 \frac{\omega}{2}} S_y(\omega)\, d\omega \ ,$$

$$R_3 = \frac{1}{mk}(1 - \frac{r}{m+r}) \int_{-\pi}^{\pi} \frac{\sin(mk\omega + \frac{rk\omega}{2})\sin\frac{rk\omega}{2}}{\sin^2 \frac{\omega}{2}} S_y(\omega)\, d\omega \ ,$$

and

$$R_4 = \frac{2}{mk} \int_{\frac{\pi}{k}}^{\pi} \frac{\sin^2 \frac{mk\omega}{2}}{\sin^2 \frac{\omega}{2}} S_y(\omega)\, d\omega \ .$$

Fix $\varepsilon > 0$. Then

$$|R_1| \leq \frac{k}{2} \int_0^{\delta} \frac{\omega^2}{\sin^2 \frac{\omega}{2}} S_k(\omega)\, d\omega + \frac{2k}{m^2} \int_{\delta}^{\pi} \frac{S_k(\omega)}{\sin^2 \frac{\omega}{2}} d\omega < \varepsilon$$

by first choosing δ so small that the first intergal is less than $\varepsilon/2$ and then choosing m so large that the same is true for the second integral. Thus $R_1 \to 0$ as $n \to \infty$. By similar arguments it follows that R_2, R_3, and R_4 also converge to 0 as $n \to \infty$. By Fejér's theorem (p.89 of Zygmund[27]), as $n \to \infty$ the integral in equation (27) converges to

$$L = \frac{1}{2} \{ f(0+) + f(0-) \} = L_1 + L_2 \ ,$$

where

$$L_1 = \pi \sum_{l=1}^{k-1} \{ S_y(\frac{2\pi l}{k}+) + S_y(\frac{2\pi l}{k}-) \}$$

and

$$L_2 = \lim_{\omega \to 0} 2\pi \{ S_y(\frac{\omega}{k})[1 - \frac{\sin^2 \frac{\omega}{2}}{k^2 \sin^2 \frac{\omega}{2k}}]\}$$

from the fact that S_y is an even function. Since $\{y_t\}$ is a power law l.m.p., $S_y(\omega) = O(|\omega|^\alpha)$ as $\omega \to 0$ by lemma 16. It is easy to show that

$$\frac{\sin^2 \frac{\omega}{2}}{k^2 \sin^2 \frac{\omega}{2k}} = 1 + O(\omega^2),$$

from which we may deduce that

$$L_2 = \lim_{\omega \to 0} O(|\omega|^\alpha) \cdot O(\omega^2) = 0.$$

The lemma now follows. $\square$

Theorem 28: Suppose that $\{y_t\}$ is a power law l.m.p. of order $-1 < \alpha < 0$ with s.d.f. S_y such that the limits $S_y(\omega+)$ and $S_y(\omega-)$ exist for all $\omega \in (-\pi, \pi)$. Then $e(k) = 1$ for all k.
Proof: By lemma 25, for any k

$$n\{var \ \bar{y}_n(k) - var \ \bar{y}_n\} = L_1 + o(1).$$

We may rewrite the above as

$$\frac{var \ \bar{y}_n(k)}{var \ \bar{y}_n} = 1 + \frac{L_1}{n \cdot var \ \bar{y}_n} + o(\frac{1}{n \cdot var \ \bar{y}_n}).$$

Since

$$n \cdot var \ \bar{y}_n = \frac{1}{n} \int_{-\pi}^{\pi} \frac{\sin^2 \frac{n\omega}{2}}{\sin^2 \frac{\omega}{2}} S_y(\omega) \ d\omega$$

and $S_y(0+) = S_y(0-) = \infty$, by the version of Fejér's theorem given by Zygmund[27], $n \cdot var \ \bar{y}_n \to \infty$. Hence

$$\frac{1}{e(k)} = \lim_{n \to \infty} \frac{var \ \bar{y}_n(k)}{var \ \bar{y}_n} = 1$$

as claimed. $\square$

The contrapositive of the next lemma shows that the conclusion of theorem 28 can only hold for stationary processes that are l.m.p.'s in the covariance sense.

Lemma 29: Let $\{y_t\}$ be a stationary process with a.c.f. C_y. If $\sum |C_y(\tau)| < \infty$, there exists a k_0 such that $e(k) < 1$ for all $k \geq k_0$.
Proof: If $\sum |C_y(\tau)| < \infty$, then $\{y_t\}$ has a continuous s.d.f. S_y by theorem 3.1.9, p.110, of Fuller[8]. Likewise, the k-decimated mean has a continuous s.d.f. S_k given by equation (26). Since both S_y and S_k are continuous, it follows from theorem 6.1.2, p.232, of Fuller[8] that as $n \to \infty$

$$var \ \bar{y}_n = \frac{2\pi S_y(0)}{n} + o(1)$$

and

$$var \ \bar{y}_n(k) = \frac{2\pi k S_k(0)}{n} + o(1).$$

Since S_y is continuous and strictly positive on at least one interval of non-zero length, it follows that

$$kS_k(0) = \sum_{l=0}^{k-1} S_y(\frac{2\pi l}{k}) > S_y(0)$$

for all $k \geq k_0$ for some positive integer k_0. Thus

$$e(k) = \frac{S_y(0)}{kS_k(0)} < 1$$

for all $k \geq k_0$ and the lemma is proved. $\square$

In the case where the a.r.e. between two competing estimators for a quantity is 1, Hodges and Lehmann[13] propose the use of deficiency as a useful criterion for distinguishing between the estimators. Given $x_1, \cdots, x_n$, consider the estimator $\bar{x}_n$ of μ_x with variance $var \ \bar{x}_n$. If it exists, let $d = d(k;n)$ be the smallest integer such that

$$var \ \bar{x}_{n+d}(k) \leq var \ \bar{x}_n < var \ \bar{x}_{n+d-1}(k). \tag{30}$$

Then $d(k;n)$ is called the deficiency of $\bar{x}_n(k)$ with respect to $\bar{x}_n$. Typically it represents the number of additional observations needed by a k-decimated mean to perform approximately as well as the sample mean with n observations. If it exists,

$$d(k) \equiv \lim_{n \to \infty} d(k;n)$$

is called the asymptotic deficiency.

Theorem 31: For a f.d.p. with $-1 < \alpha < 0$, $d(k) = \infty$ for all $k \geq 2$.

Proof: The first portion of our argument is a slight modification of one due to Hodges and Lehmann[13]. It is shown in corollary 38 below that $var \ \bar{x}_n$ is a strictly decreasing function of n. It follows from the proof of that same corollary that $var \ \bar{x}_n(k)$ is a decreasing function of n for $n \geq k$. Since $var \ \bar{x}_1 = var \ \bar{x}_k(k)$; since $var \ \bar{x}_n > 0$ for all n; and since both $var \ \bar{x}_n$ and $var \ \bar{x}_n(k)$ decrease to 0 as $n \to \infty$ by lemma 19, it follows that for $n \geq k$ there is a unique integer $d = d(k;n)$ such that equation (30) holds. Since $var \ \bar{x}_n \to 0$ as $n \to \infty$, it follows from that equation that $var \ \bar{x}_{n+d}(k) \to 0$ also. By lemma 19,

$$var \ \bar{x}_{n+d}(k) = \frac{c}{(n+d)^{1+\alpha}} + \frac{b_n}{n+d} + o(\frac{1}{n+d}) \tag{32}$$

where $c = 2h_\alpha/\{\alpha(\alpha-1)\}$ and $\{b_n\}$ is such that $|b_n| < b < \infty$ for all n and for some constant b. Thus $n+d \to \infty$. Lemma 19 further shows that

$$var \ \bar{x}_n = \frac{c}{n^{1+\alpha}} + \frac{a_n}{n} + o(\frac{1}{n}) \tag{33}$$

for $\{a_n\}$ such that $|a_n| < a < \infty$ for all n and some constant a. By equation (30) we have

$$\frac{c}{(n+d)^{1+\alpha}} + \frac{b_n}{n+d} + o(\frac{1}{n+d}) \leq \frac{c}{n^{1+\alpha}} + \frac{a_n}{n} + o(\frac{1}{n})$$

$$\leq \frac{c}{(n+d-1)^{1+\alpha}} + \frac{b_n}{n+d-1} + o(\frac{1}{n+d-1}) \ .$$

Since $n+d \to \infty$ it follows that

$$\frac{c}{n^{1+\alpha}} + \frac{a_n}{n} + o(\frac{1}{n}) = \frac{c}{(n+d)^{1+\alpha}} + \frac{b_n}{n+d} + o(\frac{1}{n+d})$$

or

$$\frac{1}{n^{1+\alpha}}\{c + \frac{a_n+o(1)}{n^{-\alpha}}\} = \frac{1}{(n+d)^{1+\alpha}}\{c + \frac{b_n+o(1)}{(n+d)^{-\alpha}}\} \ .$$

The above indicates that $(n+d)/n \to 1$ as $n \to \infty$ and hence that $d/n = o(1)$. We may now write

$$1 + \frac{d}{n} = [c + \frac{a_n+o(1)}{n^{-\alpha}}]^{-\frac{1}{1+\alpha}} [c + \frac{b_n+o(1)}{(n+d)^{-\alpha}}]^{\frac{1}{1+\alpha}}$$

$$= [1 - \frac{a_n+o(1)}{(1+\alpha)cn^{-\alpha}}] \cdot [1 + \frac{b_n+o(1)}{(1+\alpha)cn^{-\alpha}}]$$

$$= 1 + \frac{b_n-a_n+o(1)}{(1+\alpha)cn^{-\alpha}} \ .$$

Thus we have

$$d = n^{1+\alpha}\frac{b_n-a_n+o(1)}{(1+\alpha)c} \ .$$

The lemma follows if we can show that $\liminf_{n \to \infty}(b_n-a_n) > 0$ for all $k \geq 2$. By equations (32) and (33), this is the same as showing that

$$\liminf_{n \to \infty} n\{var \ \bar{x}_n(k) - var \ \bar{x}_n\} > 0 \ . \tag{34}$$

It follows from lemma 25 that

$$\lim_{n \to \infty} n\{var \ \bar{x}_n(k) - var \ \bar{x}_n\} = 2\pi\sum_{l=1}^{k-1} S_x(\frac{2\pi l}{k}) > 0 \ .$$

Thus condition (34) holds and the lemma follows. $\square$

Table 2 shows the exact deficiencies for the same values of n, α, and k as are used in table 1. The number in parentheses below each value of $d(k;n)$ represents the total number of data points that the k-decimated mean utilizes. This number is simply $\{n + d(k;n)\}/k$; the ratio of this number to n (the number of data points that the sample mean uses) approaches $1/k$ as $n \to \infty$, since $d(k;n)/n$ is $o(1)$ by the proof of theorem 31. For example, for $n = 500$, $\alpha = -.9$, and $k = 50$, the table shows that 300 additional observations are required for a 50-decimated mean to have at least as small a variance as that of an ordinary mean of 500 observations. However, a 50-decimated mean of 800 observations actually utilizes only 16 observations. The ratio $16/500 \approx 1/50$ roughly as the above theory suggests. For the sake of comparison, the column that is labelled $d_W(k;100)$ contains the corresponding deficiencies for an uncorrelated (white noise) process.

The above results on relative efficiency and deficiency have some practical consequences. If one is interested in estimating μ_x in a given time frame from a finite

<table>
<tr><td colspan="6">**Table 2**

Deficiency of $\bar{x}_n(k)$ with Respect to $\bar{x}_n$</td></tr>
</table>

		$\alpha = -.5$		$\alpha = -.9$	
k	$d_W(k;100)$	$d(k;100)$	$d(k;500)$	$d(k;100)$	$d(k;500)$
2	100	14	30	4	4
	(100)	(57)	(265)	(52)	(252)
5	400	60	135	15	20
	(100)	(32)	(127)	(23)	(104)
10	900	130	310	40	50
	(100)	(23)	(81)	(14)	(55)
20	1900	280	660	80	100
	(100)	(19)	(58)	(9)	(30)
50	4900	700	1600	200	300
	(100)	(16)	(42)	(6)	(16)

sample of data from a f.d.p. and each observation costs a certain amount, table 1 on relative efficiencies shows that it is possible by means of decimation to reduce the cost considerably at the expense of a modest decrease in efficiency if α is close enough to -1. Conversely, if only a fixed number of observations can be made and the time frame is flexible, table 2 on deficiencies suggest that the data be spaced as far apart in time as it is practical. These comments assume that the variance is an adequate criterion for choosing an estimator.

We conclude this section by proving a lemma concerning the variance of the mean of a stationary process, a corollary of which we used above in theorem 31.

Lemma 35: Let $y_1, y_2, \cdots$ represent a wide sense stationary process with a.c.f. given by C_y such that $C_y(\tau) < C_y(0)$ for all $\tau \geq 1$. Let $\bar{y}_n$ represent the sample mean of the first n observations. Then $var\ \bar{y}_n < var\ \bar{y}_{n-1}$ if and only if

$$\sum_{\tau=1}^{n-1}(\tau - M_n)C_y(\tau) < \frac{M_n C_y(0)}{2} , \qquad (36)$$

where

$$M_n = \frac{n(n-1)}{2n-1}.$$

Proof: Without loss of generality, we may assume for convenience that $Ey_t = 0$. Consider the following minimization problem: subject to $(n-1)\alpha + \beta = 1$, find α (and hence β) such that

$$V(\alpha) = var\ \{\alpha \sum_{t=1}^{n-1} y_t + \beta y_n\}$$

is minimized. Note that $V(\frac{1}{n}) = var\ \bar{y}_n$ and $V(\frac{1}{n-1}) = var\ \bar{y}_{n-1}$. It is clear that $V(\alpha) = A\alpha^2 + B\alpha + C$ for some $A, B,$ and C and hence that $V(\alpha)$ is minimized when $\alpha = \alpha_{MIN} = -\frac{B}{2A}$. Since V is quadratic in α, an easy geometric argument shows that $var\ \bar{y}_n$ is

$<$, $=$, or $>$ *var* $\bar{y}_{n-1}$ as α_{MIN} is $<$, $=$, or $> \frac{1}{2}(\frac{1}{n} + \frac{1}{n-1})$. Some algebra shows that

$$A = E\{\sum_{t=1}^{n-1} y_t - (n-1)y_n\}^2$$

$$= E\{(\sum_{t=1}^{n-1} y_t)^2 - 2(n-1)\sum_{t=1}^{n-1} y_t y_n + (n-1)^2 y_n^2\}$$

$$= E(\sum_{t=1}^{n-1} y_t)^2 - 2(n-1)\sum_{\tau=1}^{n-1} C_y(\tau) + (n-1)^2 C_y(0)$$

and

$$B = 2E\{\sum_{t=1}^{n-1} y_t y_n - (n-1)y_n^2\}$$

$$= 2\{\sum_{\tau=1}^{n-1} C_y(\tau) - (n-1)C_y(0)\}.$$

Some algebra shows that

$$E(\sum_{t=1}^{n-1} y_t)^2 = (n-1)C_y(0) + 2(n-1)\sum_{\tau=1}^{n-1} C_y(\tau) - 2\sum_{\tau=1}^{n-1} \tau C_y(\tau).$$

We then have

$$\alpha_{MIN} \approx \frac{(n-1)C_y(0) - \sum_{\tau=1}^{n-1} C_y(\tau)}{n(n-1)C_y(0) - 2\sum_{\tau=1}^{n-1} \tau C_y(\tau)}.$$

If we substitute this equality into

$$\alpha_{MIN} < \frac{2n-1}{2n(n-1)},$$

which is true if and only if *var* $\bar{y}_n$ < *var* $\bar{y}_{n-1}$, then inequality (36) follows and the lemma is proved. □

Lemma 37: If $C_y(0) > C_y(1) > \cdots > C_y(n-1) > 0$, then inequality (36) holds.

Proof: We use an inequality of Chebyshev: if $a_1 \le a_2 \le \cdots \le a_n$ and $b_1 \ge b_2 \ge \cdots \ge b_n$, then

$$\sum_{k=1}^{n} a_k b_k \le \frac{1}{n}(\sum_{k=1}^{n} a_k)(\sum_{k=1}^{n} b_k).$$

(See 3.2.7, p. 11 of Abramowitz and Stegun[1]). We now have

$$\sum_{\tau=1}^{n-1} (\tau - M_n)C_y(\tau) \le \frac{1}{n-1}(\sum_{\tau=1}^{n-1} [\tau - M_n])(\sum_{\tau=1}^{n-1} C_y(\tau))$$

$$< C_y(0)\sum_{\tau=1}^{n-1} (\tau - M_n) = \frac{M_n C_y(0)}{2},$$

and the lemma holds. □

Corollary 38: If $\{x_t\}$ is a f.d.p. of order $-1 < \alpha < 0$, then $var\ \bar{x}_n$ is a strictly decreasing sequence.

Proof: By equation (14)

$$C_x(\tau) = \frac{2\tau-\alpha-2}{2\tau+\alpha}C_x(\tau-1) < C_x(\tau-1)$$

for all $\tau \geq 1$. This corollary is thus an immediate consequence of lemma 37. $\square$

Let us note in passing that $var\ \bar{y}_n \leq var\ \bar{y}_{n-1}$ need not hold for all n for all stationary processes. Let $\{y_t\}$ be a second order autoregressive process with $\phi_1 = 0$ and $|\phi_2| < 1$. (By pages 60-61 of Box and Jenkins[4], these coefficients do yield a stationary process.) Now $\rho_1 = 0$ and $\rho_2 = \phi_2$, and it is easy to show that $var\ \bar{y}_3 > var\ \bar{y}_2$ if $\rho_2 > .75$.

1.4 The Sample Variance of a F.D.P.

Let $\{x_t\}$ represent a f.d.p. of order $-1 < \alpha < 0$ with a.c.f. C_x. Let us assume that both $Ex_t = \mu$ and $var\ x_t = C_x(0)$ are unknown and that we want to estimate $C_x(0)$ from a data sample $x_1, x_2, \cdots, x_n$. A natural estimator to consider is the sample variance

$$s_n^2 = \frac{1}{n}\sum_{t=1}^{n}(x_t - \bar{x}_n)^2.$$

It can be shown (see David[6]) that for any stationary process

$$0 \leq Es_n^2 \leq C_x(0).$$

The following lemma shows that this estimator can be severely biased for a f.d.p.

Theorem 39: For every sample size N and every $\varepsilon > 0$, there exists a f.d.p. of order $-1 < \alpha < 0$ such that

$$Es_n^2 < \varepsilon C_x(0) \tag{40}$$

for all $n \leq N$.

Proof: Fix $\varepsilon > 0$ and N. First, note that

$$Es_n^2 = \frac{1}{n}\sum_{t=1}^{n}E\{(x_t - \mu) - (\bar{x}_n - \mu)\}^2$$

$$= C_x(0) - var\ \bar{x}_n. \tag{41}$$

By the use of equation (13) and the second part of lemma 19, the above may be rewritten as

$$\frac{Es_n^2}{C_x(0)} = 1 - \frac{\Gamma(1+\frac{\alpha}{2})}{n^{1+\alpha}(1-\alpha)\Gamma(1-\frac{\alpha}{2})} + O(\frac{1}{n}) + O(\frac{1}{(1+\alpha)n^2}),$$

where the bounds for the O terms may be chosen independently of α. Let $\alpha = -1 + \frac{1}{n}$. We then have

$$\frac{Es_n^2}{C_x(0)} = 1 - \frac{\Gamma(\frac{1}{2}+\frac{1}{2n})}{n^{\frac{1}{n}}(2-\frac{1}{n})(\frac{1}{2}-\frac{1}{2n})\Gamma(\frac{1}{2}-\frac{1}{2n})} + O(\frac{1}{n}).$$

Since the term with the gamma functions in the above equation converges to 1 as $n \to \infty$, it is possible to find a n_0 such that equation (40) holds for all $n \geq n_0$. Let n_1 be the greater of n_0 and N and set $\alpha_1 = -1 + \dfrac{1}{n_1}$. Then equation (40) holds when $n = n_1$ for a f.d.p. of order α_1. We need only show that it also holds when $n < n_1$ to complete the lemma. By corollary 38, $var \; \bar{x}_n$ is a strictly decreasing function of n, which means that Es_n^2 is a strictly increasing function of n because of equation (41). Hence the left hand side of equation (40) is also a strictly increasing function of n, which shows that (40) holds for all $n \leq n_1$. $\square$

The thrust of this theorem is that s_n^2 can severely underestimate $C_x(0)$ even for very large sample sizes if α is close to -1, a fact that Allan[2] noted in a slightly different form in his 1966 paper. Table 3 gives values of the ratio of Es_n^2 to $C_x(0)$ for various α's and sample sizes n. For contrast, the column that is marked "white" gives the corresponding values for a white noise process.

Table 3

Ratio of Es_n^2 to the Process Variance

n	white	α -.50	-.75	-.90	-.95	-.99
50	.9800	.87	.65	.35	.20	.04
100	.9900	.91	.71	.39	.22	.05
500	.9980	.96	.80	.48	.28	.06
1000	.9990	.97	.84	.52	.31	.07
5000	.9998	.99	.89	.59	.36	.09
10000	.9999	.99	.91	.62	.38	.09

One hope for correcting this bias in s_n^2 is to estimate α by some means. Mohr[16] utilizes this procedure in her study of the Nile River data for which $\alpha \approx -.7$. Table 3 shows, however, that the bias depends strongly on α for α close to -1, so attempts to produce reasonable estimates of $C_x(0)$ for such processes by a bias correction of s_n^2 seem hopeless.

References

1. M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions*, U.S. Government Printing Office, Washington, D. C., Dec. 1972. (tenth printing)

2. D. W. Allan, "Statistics of Atomic Frequency Standards," *Proc. IEEE*, vol. 54, no. 2, pp. 221-230, Feb. 1966.

3. T. W. Anderson, *The Statistical Analysis of Time Series*, John Wiley & Sons, New York, 1971.

4. G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, Inc., San Francisco, 1976. (revised edition)

5. D. R. Brillinger, *Time Series Data Analysis and Theory*, Holden-Day, Inc., San Francisco, 1981. (expanded edition)

6. H. A. David, "Bias of S^2 Under Dependence," *American Statistician*, vol. 39, p. 201, 1985.

7. J. E. Deeter, P. E. Boynton, and D. B. Percival, "Vela X-1 Timing Experiment," Progress Report, Department of Astronomy, University of Washington, Seattle, Oct. 1981.

8. W. A. Fuller, *Introduction to Statistical Time Series*, John Wiley & Sons, New York, 1976.

9. C. W. J. Granger, "The Typical Spectral Shape of an Economic Variable," *Econometrica*, vol. 34, no. 1, pp. 150-161, Jan. 1966.

10. C. W. J. Granger and R. Joyeux, "An Introduction to Long-memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, vol. 1, pp. 15-29, 1980.

11. E. J. Hannan, *Multiple Time Series*, John Wiley & Sons, New York, 1970.

12. F. B. Hildebrand, *Introduction to Numerical Analysis*, McGraw-Hill Book Company, New York, 1974. (second edition)

13. J. L. Hodges and E. L. Lehmann, "Deficiency," *Annals of Math. Stat.*, vol. 41, no. 3, pp. 783-801, 1970.

14. A. V. Holden, *Models of the Stochastic Activity of Neurones*, Springer-Verlag, Berlin, 1976.

15. B. B. Mandelbrot, "Statistical Methodology for Nonperiodic Cycles: From the Covariance to R/S Analysis," *Annals of Economic and Social Measurement*, vol. 1, no. 3, pp. 259-290, July 1972.

16. D. L. Mohr, "Estimating the Parameters of a Fractional Gaussian Noise," Technical Report, Department of Mathematics, Tulane University, New Orleans, 1981.

17. W. H. Munk and G. J. F. MacDonald, *The Rotation of the Earth*, Cambridge University Press, Cambridge, 1975.

18. T. Musha and H. Higuchi, "The 1/f Fluctuation of a Traffic Current on an Expressway," *Japanese Journal of Applied Physics*, vol. 15, no. 7, pp. 1271-1275, July 1976.

19. E. Parzen, "Time Series Model Identification and Prediction Variance Horizon," in *Applied Time Series Analysis II*, ed. D. F. Findley, pp. 415-447, Academic Press, Inc.,

New York, 1981.

20. D. B. Percival, "Prediction Error Analysis of Atomic Frequency Standards," *Proc. 31st Annual Symposium on Frequency Control*, pp. 319-326, Atlantic City, N. J., June 1977.

21. D. B. Percival, "The Statistics of Long Memeory Processes," unpublished Ph.D. thesis, Department of Statistics, University of Washington, Seattle, 1983.

22. M. B. Priestley, *Spectral Analysis and Time Series,* Academic Press, New York, 1981.

23. K. L. Schick and A. A. Verveen, "1/f Noise with a Low Frequency White Noise Limit," *Nature*, vol. 251, pp. 599-601, 18 Oct. 1974.

24. E. C. Titchmarsh, *The Theory of Functions,* Oxford University Press, Oxford, 1979. (second edition)

25. R. F. Voss and J. Clarke, "'1/f Noise' in Music and Speech," *Nature*, vol. 258, pp. 317-318, 27 Nov. 1975.

26. R. F. Voss, "1/f (Flicker) Noise: A Brief Review," *Proc. 33rd Annual Symposium on Frequency Control*, pp. 40-46, Atlantic City, N. J., 30 May - 1 June 1979.

27. A. Zygmund, *Trigonometric Series,* Cambridge University Press, London, 1978.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
69	ADA176221	

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
On the sample mean and variance of a long memory process	TR 12/1/85-11/30/85
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
Donald B. Percival	N00014-84-C-0169

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Department of Statistics, GN-22 University of Washington Seattle, WA 98195	NR-661-003

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
ONR Code N63374 1107 NE 45th St. Seattle, WA 98105	September, 1985
	13. NUMBER OF PAGES
	22

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Sample mean; sample variance; long memory process

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

DD FORM 1473   EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

S/N 0102-LF-014-6601

END
3-87
DTIC